

Back in 2014, Elon Musk referred to AI as [summoning the demon](#). And it wasn't hard to see that view. Soon, [Go agents](#) would beat top humans learning from self play. By the end of 2017, the [same algorithm](#) mastered Chess and Shogi. By 2020, it didn't even need tons of calls to the simulator, and could [play Atari too](#).

AI looked scary. It looked like it was one FOOM away from self playing and becoming superhuman at the universe. And yet, here we are in 2023 and self driving cars still don't work, never mind your robotic maid and chef.

Does becoming superhuman at the universe make any practical sense? The universe has so many orders of magnitude more states than any Go game. And I don't even need math to illustrate this point, just imagine tiling the universe with as many very tiny Go boards as you can fit.

That's a lot of Go boards. So many that the difference in complexity between Go and the Universe is not a matter of number, it's a matter of kind. Every Go program operates at the stone level. Almost nothing predicting the world operates at the atom level.

---

These modern self play systems like [MuZero](#) have some form of dynamics model, a function that given the current state of the world and an action, it predicts the next state. We also use these dynamics models [at comma](#). They are world models, and contain all the knowledge of how the world works inside of them.

GPT-4 is a dynamics model also, conditioned on the prior of the action space. And there's an even simpler way to think about it. The loss function for dynamics is compression.

[Compression is prediction is intelligence, intelligence is prediction is compression](#). One of the coolest facts I ever learned. So, feed in the whole internet, build a compressive model, make it really really big, and you just won the universe?

---

Not so fast. This approach can lead to [very neat applications](#), but does it take over the world?

You want to get rich? Here's an idea. Download all the historical stock market data. Get lots and lots of GPUs to train a huge model. Boom, predictor for the stock market. I look at prediction, I know which stocks go up! I buy the stocks that go up, I short the stocks that go down. With mega leverage, I am a billionaire by the end of the week! I can reinvest my billions in more GPUs for recursive self improvement. Nobody will stop me, I will take over the whole economy. How is nobody doing this?!?

Oh wait...every hedge fund bro is already doing this. And most of them aren't billionaires. The problem is your model needs to include all the computers playing the market, and it also needs to include the other hedge fund bros themselves. This strategy only dominates if you have more compute than the whole market itself, which you don't.

---

In Go, your model doesn't need to include other computers. Yes, you are playing against a computer that you are modeling, but the game itself is way too small to contain a Go playing computer. The other computer you are playing against is outside the universe. The rules of Go don't include computers. The rules of Go are fixed in complexity.

On the other hand, your dynamics model of the universe must include computers. You'll have to spend a lot of effort modeling them. Unless you have an absolutely staggering advantage, you aren't going to figure them out from self play. The computers are active players, and they have a lot of compute. They might even be using it to try to model you!

Assuming [no untoward market intervention](#), why would any one system ever have a large majority of the compute? Compute will be distributed in a [power law](#).

The smart regulation isn't capping the FLOPS in training runs. That's creating a powder keg. If the FLOPS are artificially restricted, and one person breaks the restriction, you could end up with a single dominant system.

If you don't want FOOM, you just need to prevent a 51% attack on compute.

---

Now, if there's one weird trick to 1e20x your efficiency, and only one group gets it, all bets are off. But this is never how things happen. Nobody has a 1e20x more efficient steam engine, it isn't even possible.

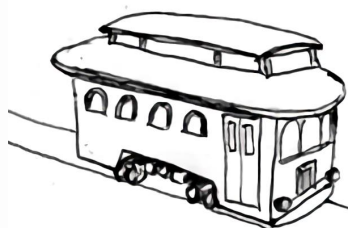
Nobody has a 1e20x more efficient Bitcoin miner either, and I also doubt that's possible, we just [understand](#) steam engines a lot better, so for steam engines we *know*. More intelligence leads to more new tricks, but the tricks get harder and harder to find.

There's low hanging fruit, you pick it, then you build tools to get the higher hanging fruit. You spend money on ladders, electric crane thingies, more and more money to get less and less fruit. Oil [used to be](#) just pouring out of the ground, now we go [to the bottom](#) of the ocean.

---

A revolution is coming. The information revolution will do for intelligence what the industrial revolution did for energy. Most work (force\*distance) used to be done by muscles, now it's not. Most thinking used to be done by brains, soon it won't be.

But it's not going to happen overnight. It'll happen on a nice exponential, like it already is. The universe is an unfathomable number of orders of magnitude more complex than the Go game. The universe includes the player. The universe includes the other players. Games don't.



**There is no problem.  
You are travelling safely to  
your destination.**

Unless we build a terrifying powder keg, there is no FOOM. Let the markets cook.